

# Revisiting Morals Clauses In The Age Of Deepfakes

By **Dyan Finguerra-DuCharme, Simon Pulman and Laure Sawaya** (June 21, 2024)

Taylor Swift. Selena Gomez. Piers Morgan. Oprah Winfrey. All of these public figures have been implicated in misleading deepfake content created with artificial intelligence in the past 12 months.

Because reputation is paramount for talent, companies that contract with celebrities, actors, singers or other public figures are engaging those individuals' services — and buying the goodwill that comes with their reputation, personal brand and following. Accordingly, if the talent's reputation becomes tarnished by a scandal or public controversy, both parties are harmed, and the value of the company's investment in the talent is diminished.

Even before the threat of AI began to rise, reputational due diligence and contractual protections have been critical when entering into talent agreements. At the outset of the relationship, the emphasis is on the company to research and vet the talent's background and reputation.

This process continues throughout the duration of the relationship to guard against reputational harm, which requires vigilance and, at times, close communication between the company and the talent's team to navigate fraught situations and potential misinformation.

The growing sophistication and ubiquity of AI technology has complicated the traditional process of reputation management, and is poised to alter the relationship between contracting parties exponentially in years to come.

Software such as Midjourney and Sora can allow users to create relatively believable images and videos using text prompts, and this has already led to a proliferation of unauthorized content using the name, likeness and other identifying characteristics of high-profile talent. At its worst, this content may paint individuals in an unsavory light, jeopardizing their contractual relationships and causing lasting reputational damage.

As noted above, we have seen deepfake scandals in the past year that have implicated multiple high-profile public figures, including consumer scams and unauthorized deepfake pornography. Earlier this year, a Channel 4 News analysis found that over 4,000 famous people have had their faces superimposed on pornographic material through the use of AI technology.[1] That number is only likely to climb.

To shield against the general risk of reputational damage, talent agreements typically include a morals or conduct clause. These clauses have been commonplace in endorsement and sponsorship agreements for years, and have recently become commonplace in all forms of film and television agreements in the aftermath of the #MeToo era.

The purpose of a morals clause is to provide an express remedy against talent whose



Dyan  
Finguerra-DuCharme



Simon Pulman



Laure Sawaya

conduct, statements and/or associations adversely affect or are likely to adversely affect the public's perception of that talent, and therefore any project, brand or company with whom the talent is associated.

In other words, the reputation of the talent is highly likely to have an impact on the potential success of the motion picture, television series, marketing campaign, product or other endeavor involving them, thus affecting the value of the accompanying investment and the goodwill associated with a corporate brand.

The digital era has complicated reputation management in several ways. Social media created a means for celebrities to interface directly with their audience, but also presented new risks for them, such as potential controversial statements. There have been several high-profile examples of old social media posts being resurfaced and pulling individuals into present and very real controversies — in some instances causing them to lose work in the process.

AI has made it easier to search the internet's archives of current and historic files on people or events that could allow the public to uncover and disseminate information about public figures. It will also accelerate the creation and dissemination of false information about celebrities that could give rise to legal action by or against the talent. For example, AI tools can make it much easier for false information to be disseminated using tools that create fake articles and other information that can have an uncanny resemblance to a real publication.

The deepfake — or, in the parlance of the Screen Actors Guild-American Federation of Television and Radio Artists, "digital replica" — is one of the most prevalent and powerful forms of disinformation affecting celebrities. These AI-generated images and videos of individuals are designed to give the impression of a real photograph or video recording of the individual. While deepfakes can be made of anyone, some of their most prevalent and insidious uses involve celebrities and public figures.

For example, it is relatively easy for an internet user with access to AI tools to create photos or videos of a celebrity engaged in activity that would be objectionable to that celebrity and to their creative and commercial partners, such as endorsing a rival brand, expressing a statement that may be political or offensive, or endorsing a scam.

This use of deepfakes creates a potential complication as to what constitutes — or, rather, should constitute — "conduct" to ensure that the purpose of the morals clause is met. For the first time, it is possible for bad actors to create extremely authentic-seeming photographs, videos, and audio or voice recordings in a matter of seconds using AI — and to easily disseminate misleading or misrepresentative content without the talent having actually engaged in any "conduct" per se.

Take, for example, videos that were circulated earlier this year, in which deepfaked replicas of Taylor Swift and Selena Gomez announced to fans that there is an excess in inventory for Le Creuset Dutch ovens, encouraging fans to act quickly to claim one by clicking a button and providing certain information, including a credit card for the shipping cost.

Numerous fans believed the videos to be authentic and bought into the scam, providing personal information to the scammers. When they did not receive the product, they blamed Le Creuset, thereby damaging the company's goodwill, among other consequences.

There is also the recent example of a high school athletics director in Baltimore, Dazhon

Darien, who was arrested in April for allegedly using AI to fake the voice of the school's principal engaged in offensive hate speech. The clip went viral in January.[2]

AI technology is already being used in many criminal activities and is likely to give rise to new forms of consumer fraud. While some of that may include deepfake likenesses and voices of noncelebrities, some of it will surely include public figures.

What does this mean for morals clauses? Here are some likely scenarios.

### **A Challenge for Talent**

It seems inevitable that famous talent will be subject to some kind of AI misrepresentation — ranging from the humorous and relatively benign, to the sensitive and intrusive, to the downright offensive and reputation-threatening. Some of these deepfakes could certainly fall within the sphere of subjecting the talent and their employers to public disrepute, contempt, scandal, or ridicule, which is usually the kind of impact a morals clause shields against.

In light of this, we can expect to see talent push for ever-tighter legislation and regulation (state and federal, as well as at the guild level) to protect them. Contractually, it seems likely that talent reps will seek express language to ensure that their clients are not unfairly penalized for misuse of AI by third-party bad actors.

### **A Challenge for Companies Engaging Talent**

Companies will likely also be thinking more expansively about how a morals clause is drafted and what events should be anticipated. Because reputational harm is the cornerstone of the morals clause, and such harm can now be created without any actual conduct of the talent, we expect to see a push for broader and wider-reaching morals clauses that will cover public accusations and perceptions of misconduct as opposed to mere misconduct.

We can also expect to see companies trying to expand the time frame during which the accusations or alleged conduct took place. AI deepfakes can be made to appear to have taken place in the far past but be disseminated, and the fabricated "conduct" allegedly "discovered," only today.

Similarly, the power of AI as a search tool might also change how easily one might uncover truthful past misconduct otherwise buried on the internet. This might help companies better vet the talent with whom they work, or can provide another consideration about past conduct to be addressed in the morals clause.

### **A Complication for Investigations**

An alleged violation of a morals clause, particularly when involving discrimination, harassment or sexual misconduct, is usually followed by an investigation. AI threatens to complicate those investigations.

Take the example of a celebrity who is alleged to have said something discriminatory or threatening that was allegedly recorded by a third party. An obvious defense is that the talent did not make the alleged statement and that the recording was synthesized via AI — easy to do for any public figure with hours or even just minutes of easily accessible performances and interviews.

Accordingly, investigators will need to familiarize themselves with AI technology, and probably partner with technology companies that can detect AI usage, or lack thereof, by technical analysis and use of watermarks. Expect this to become a reasonably routine part of investigations and, potentially, any related litigation.

### **Setting the Record Straight**

Even if it is proven that damaging content was generated by AI, the damage to a brand, company or production may already be done — especially given how quickly misinformation can spread on the internet. Accordingly, we are starting to see examples of brands requiring that the talent in influencer, sponsorship and endorsement contracts take reasonable remedial measures to correct the misinformation via social media channels or elsewhere.

Typically, we can expect to see an obligation that the talent actively engage in public discussions about the misinformation and that they reasonably cooperate with the company in undertaking any remedial campaign necessary. The major question becomes: At whose cost should such additional services be performed?

Unlike a typical breach of contract situation where the burden of curing misconduct and all costs associated therewith can equitably be imposed on the party who engaged in that behavior, where the harm is caused by a third party, ideas of fairness would call for a shared burden on the company and the talent.

The advent of AI has brought new challenges for public figures and the companies engaging them. Both must recognize that false and deceptive information about or involving celebrities that can harm their reputation has become not just possible, but rather incredibly easy.

As a result, it is now more important than ever for parties to an agreement involving the use of someone's persona to anticipate and guard against such concerns and preemptively establish remedial protocols to address such issues in the unfortunate event that they occur.

---

*Dyan Finguerra-DuCharme is a partner and co-chair of the intellectual property group at Pryor Cashman LLP.*

*Simon Pulman is a partner and co-chair of the media and entertainment, and film, TV and podcast groups at the firm.*

*Laure Sawaya is special counsel at the firm.*

*The opinions expressed are those of the author(s) and do not necessarily reflect the views of their employer, its clients, or Portfolio Media Inc., or any of its or their respective affiliates. This article is for general information purposes and is not intended to be and should not be taken as legal advice.*

[1] <https://www.theguardian.com/technology/2024/mar/21/celebrities-victims-of-deepfake-pornography>.

[2] [www.nytimes.com/2024/04/25/technology/deepfake-recording-principal-arrest.html](http://www.nytimes.com/2024/04/25/technology/deepfake-recording-principal-arrest.html).